**Worksheet 4.1B- Correlation NEED: Program CORR, group HYPOCORR**

* Ungroup HYPOCORR (you should get lists EXA1, EXA2, EXB1, EXB2, etc.)
* These lists are hypothetical exam scores, the same ones from the 4.1A worksheet

1. Look back at the scatterplots from the section 4.1A worksheet to see the graphs for Classes A-F. Now compute the **correlation** for each graph, and fill in the correlation in the same table from before:

(EXA1 = x-list, EXA2 = y-list ….. EXB1 = x-list, EXB2 = y-list…. etc.)

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Strong** | **Moderate** | **Weak** |
| **Negative** | C | D | F |
| **Positive** | E | A | B |

1. Based on these results, what do you suspect is the largest value that a correlation can assume? What do you suspect is the smallest value?
2. What scatterplots have the largest correlation? Which have the smallest?
3. How does the value of the correlation relate to the *direction* of the association?
4. How does the value of the correlation relate to the *strength* of the association?
5. Make a scatterplot relating the scores from Exam G (EXG1 = x-list, EXG2 = y-list). Draw the graph. Describe the scatterplot.
6. Calculate the correlation for Class G. What type of relationship does *r* REALLY measure??
7. Calculate the correlation for Class G but switch the X and Y list. What is it? How does it compare to the original correlation? Do the same for Class A (switch X and Y). What does this tell you?
8. Make a scatterplot for Class H (EXH1 = x-list, EXH2 = y-list). Sketch it. Do most of the observations seem to follow a linear pattern? Are there any exceptions?
9. Make a scatterplot for Class I (EXI1 = x-list, EXI2 = y-list). Sketch it. Do most of the observations seem to be scattered? Are there any exceptions?
10. Calculate the correlation for Classes H and I. Write them both down. Does either of these correlations surprise you? Why?
11. Remove the outlier for Class H (be sure to remove the coordinates from both the X and Y lists). Recalculate the correlation, and write it down. How has it changed?
12. Remove the outlier for Class I. Recalculate the correlation, and write it down. How has it changed?
13. Based on your analysis of Classes H and I, would you say that the correlation coefficient is affected by outliers?
14. Make a scatterplot for Class J. Sketch it. Describe what the plot reveals about the relationship between exam scores (describe the plot).
15. Calculate the correlation for Class J. Does this value surprise you? How?

**Worksheet 4.1B- Correlation NEED: Program CORR, group HYPOCORR**

* Ungroup HYPOCORR (you should get lists EXA1, EXA2, EXB1, EXB2, etc.)
* These lists are hypothetical exam scores, the same ones from the 4.1A worksheet

1. Look back at the scatterplots from the section 4.1A worksheet to see the graphs for Classes A-F. Now compute the **correlation** for each graph, and fill in the correlation in the same table from before:

(EXA1 = x-list, EXA2 = y-list ….. EXB1 = x-list, EXB2 = y-list…. etc.)

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Strong** | **Moderate** | **Weak** |
| **Negative** | C | D | F |
| **Positive** | E | A | B |

1. Based on these results, what do you suspect is the largest value that a correlation can assume? What do you suspect is the smallest value?
2. What scatterplots have the largest correlation? Which have the smallest?
3. How does the value of the correlation relate to the *direction* of the association?
4. How does the value of the correlation relate to the *strength* of the association?
5. Make a scatterplot relating the scores from Exam G (EXG1 = x-list, EXG2 = y-list). Draw the graph. Describe the scatterplot.
6. Calculate the correlation for Class G. What type of relationship does *r* REALLY measure??
7. Calculate the correlation for Class G but switch the X and Y list. What is it? How does it compare to the original correlation? Do the same for Class A (switch X and Y). What does this tell you?
8. Make a scatterplot for Class H (EXH1 = x-list, EXH2 = y-list). Sketch it. Do most of the observations seem to follow a linear pattern? Are there any exceptions?
9. Make a scatterplot for Class I (EXI1 = x-list, EXI2 = y-list). Sketch it. Do most of the observations seem to be scattered? Are there any exceptions?
10. Calculate the correlation for Classes H and I. Write them both down. Does either of these correlations surprise you? Why?
11. Remove the outlier for Class H (be sure to remove the coordinates from both the X and Y lists). Recalculate the correlation, and write it down. How has it changed?
12. Remove the outlier for Class I. Recalculate the correlation, and write it down. How has it changed?
13. Based on your analysis of Classes H and I, would you say that the correlation coefficient is affected by outliers?
14. Make a scatterplot for Class J. Sketch it. Describe what the plot reveals about the relationship between exam scores (describe the plot).
15. Calculate the correlation for Class J. Does this value surprise you? How?